UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/046,131	10/21/2001	Francisco M. Galanes	M61.12-0393	9228

69316 7590 07/02/2010

MICROSOFT CORPORATION
ONE MICROSOFT WAY
REDMOND, WA 98052

EXAMINER	
LERNER, MARTIN	
ART UNIT	**PAPER NUMBER**
2626	

NOTIFICATION DATE	DELIVERY MODE
07/02/2010	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

DBOUTON@MICROSOFT.COM
vffiling@microsoft.com
stevensp@microsoft.com

Office Action Summary	Application No.	Applicant(s)	
	10/046,131	GALANES ET AL.	
	Examiner	Art Unit	
	MARTIN LERNER	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *07 May 2010*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1 and 3 to 56* is/are pending in the application.

4a) Of the above claim(s) *53 to 56* is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1 and 3 to 52* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *02 June 2006* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

Election/Restrictions

Applicants' election without traverse of Group II, Claims 23 to 51, in the reply filed on 13 June 2008 is acknowledged.

Claims 53 to 56 are withdrawn from further consideration pursuant to 37 CFR 1.142(b) as being drawn to a nonelected invention, there being no allowable generic or linking claim. Election was made **without** traverse in the reply filed on 13 June 2008.

Claim Rejections - 35 USC § 112

The following is a quotation of the first paragraph of 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

Claims 1 and 3 to 22 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the written description requirement. The claims contain subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventors, at the time the application was filed, had possession of the claimed invention.

Independent claims 1 and 12 contain the terms "modality dependent attributes" and "modality dependent controls", which are new matter because Applicants' Specification as originally-filed does not provide an adequate written description in such a way as to reasonably convey that the inventors had possession of the concept of

modality dependence at the time of original filing. The Specification does not set forth the term "modality dependent", and the only suggestion of the term is from *Dantzig et al.*, the prior art from which Applicants are attempting to distinguish. Thus, there is no express written disclosure of the term “modality dependent” from Applicants’ Specification.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 1 and 4 to 8 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Dantzig et al*. in view of *Coffman et al*.

Concerning independent claim 1, *Dantzig et al*. discloses a system and method for generating multi-modal applications from markup scripts, comprising:

“a set of controls defined in an authoring page for a website for defining visual renderings and at least one of recognition and audible prompting on a client in a client/server system, each control having a first set of attributes related to visual rendering and a second set of attributes related to at least one of recognition and audibly prompting, wherein one of the second set of attributes for one of the controls relates to a grammar to use for recognition, the controls being related to client side markup executable by a client browser” – an XML (eXtensible Markup Language) script

is implemented in a single authoring format ("an authoring page") (column 5, lines 50 to 56); main renderer 14 of a multi-modal presentation manager 11 initiates a first processing thread comprising a GUI presentation manager 15 ("a first set of attributes related to visual rendering") (column 7, lines 38 to 43: Figure 1); presentation of a graphic user interface (GUI) for an application defines a "visual rendering"; main renderer 14 of a multi-modal presentation manager 11 initiates a second processing thread comprising a speech renderer 16 ("a second set of attributes related to at least one of recognition and audibly prompting"), wherein the speech renderer 16 utilizes a plurality of speech engines 17 for speech recognition and text-to-speech synthesis (column 7, lines 38 to 47: Figure 1); controls are "modality dependent" because each processing thread is directed to either a modality relating to GUI presentation or a modality relating to a speech renderer; multi-modal presentation manager 11 controls an application on a web browser or a desktop (column 8, lines 32 to 35: Figure 1); one thread comprising a GUI presentation manager 15 is "related" to defining visual renderings on the client device because the thread initiates a visual modality; similarly, a second thread comprising a speech renderer 16 is "related" to defining desired operation on the client device because the thread initiates speech recognition or text-to-speech synthesis; a speech renderer utilizes grammars according to JSGF (Java Speech Grammar Format) for speech recognition (column 9, lines 31 to 39; column 16, lines 26 to 30); VoiceXML makes extensive use of grammars in order to optimize speech recognition functions ("wherein one of the second set of attributes for one of the controls relates to a grammar to use for recognition") (column 10, lines 38 to 40);

"a module operable on a computer, the module being configured to receive the authoring page, and wherein the module is further configured to generate, using modality dependent attributes provided from controls on the authoring page, client side markup executable by the client browser on the client in the server/client system in accordance with the controls and the attributes of the controls to perform both visual rendering and at least one of recognition and audibly prompting" – multi-modal presentation manager 11 controls an application on a web browser or a desktop (column 8, lines 32 to 35: Figure 1); implicitly, a web browser is executed on a client in a client/server architecture for receiving information from the Internet; a "single-authoring" system and method is an interaction-based programming paradigm for creating content as an intent-based markup script (column 5, line 20 to column 6, line 2; column 10, lines 24 to 28); thus, authoring for web-based presentation is on "an authoring page" at a client browser; main renderer 14 of a multi-modal presentation manager 11 initiates a first processing thread comprising a GUI presentation manager 15 (column 7, lines 38 to 43: Figure 1); presentation of a graphic user interface (GUI) for an application defines a "visual rendering"; main renderer 14 of a multi-modal presentation manager 11 initiates a second processing thread comprising a speech renderer 16, wherein the speech renderer 16 utilizes a plurality of speech engines 17 for speech recognition and text-to-speech synthesis (column 7, lines 38 to 47: Figure 1).

Concerning independent claim 1, the only elements arguably omitted by *Dantzig et al*. are that the attributes are "modality dependent". *Dantzig et al*. discloses that one thread comprising a GUI presentation manager and a second thread comprising a

speech renderer are generated from components of modality-independent IML input files rather than directly generating the visual rendering, recognition, and audible prompting. Still, *Coffman et al*. teaches a system and method for providing dialog management in a multi-modal environment, where an input/output (I/O) application program interface (API) 18 provides device abstractions and modality-dependent presentation based on an I/O modality or modalities being utilized. (Column 5, Line 59 to Column 6, Line 3: Figure 2) Multi-modal interaction dialog comprises modalities including speech (e.g., authored in VoiceXML) and visual (GUI) (e.g., hypertext markup language). (Column 4, Lines 17 to 23) An objective is to provide seamless, multi-modal access across a plurality of conversational applications and frameworks. (Column 1, Lines 49 to 60) It would have been obvious to one having ordinary skill in the art to provide modality-dependent attributes and controls related to visual rendering and recognition as taught by *Coffman et al*. in a system and method for generating and presenting multi-modal applications of *Dantzig et al*. for a purpose of providing seamless, multi-modal access across a plurality of conversational applications.

Concerning claim 4, *Dantzig et al*. discloses that controls relate to grammars for speech recognition (column 9, lines 31 to 39; column 16, lines 6 to 30).

Concerning claims 5 and 6, *Dantzig et al*. discloses that controls relate to XML (column 5, lines 50 to 56), VoiceXML (a form of XML) (Abstract), and WML (column 6, lines 56 to 62).

Concerning claims 7 and 8, *Dantzig et al*. discloses a speech renderer 16 generates audible output by text-to-speech synthesis (column 7, lines 42 to 45).

Claims 3 and 9 to 11 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Dantzig et al*. in view of *Coffman et al*. as applied to claims 1 and 2 above, and further in view of *Ladd et al. ('336)*.

Concerning claim 3, *Dantzig et al*. omits attributes for grammars and retrieving grammars from database locations. However, *Ladd et al. ('336)* teaches attributes for grammars (column 13, lines 6 to 10), and retrieving grammars from database locations (column 12, lines 7 to 14; column 14, lines 18 to 28) for speech recognition. *Ladd et al. ('336)* discloses a voice browser for interactive services, where a GRAMMAR input includes a SCR attribute that can be a grammar address (*i.e.*, a URL) for a markup language document: SCR = "gram//.SomeGrammar/month/year" ("location of a grammar for use in recognition"). (Column 20, Line 47 to Column 21, Line 1) An objective is permit users to access information from any location in the world via any suitable network access device. (Column 43, Lines 54 to 63) It would have been obvious to one having ordinary skill in the art to include markup attributes relating to a location of a grammar as taught by *Ladd et al. ('336)* in a system and method for generating and presenting multi-modal applications from markup scripts of *Dantzig et al*. for a purpose of permitting users to access information from any location in the world via a suitable network access device.

Concerning claims 9 to 11, *Ladd et al. ('336)* discloses determining an address for playing a prompt to a user (column 13, line 66 to column 14, line 17: Figure 5a:

Steps 400, 402, 406); both recorded sound samples (column 15, line 63) and text to speech (TTS) (column 16, lines 11 to 20) are provided.

Claims 12 to 46 and 52 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Dantzig et al*. in view of *Ladd et al. ('336)*.

Concerning independent claims 12, 23, and 52, *Dantzig et al*. discloses a system and method for generating multi-modal applications from markup scripts, comprising:

"a first set of visual controls having attributes related to a first modality of interaction with a user of the client that being visual renderings on the client device, the first set of controls being related to client side markup executable by a client browser" – main renderer 14 of a multi-modal presentation manager 11 initiates a first processing thread comprising a GUI presentation manager 15; an XML (eXtensible Markup Language) script is implemented in a single authoring format ("on an authoring page for a website") (column 5, lines 50 to 56); presentation of a graphic user interface (GUI) for an application defines "visual renderings"; multi-modal presentation manager 11 controls an application on a web browser or a desktop (column 8, lines 32 to 35: Figure 1); implicitly, a web browser is executed on a client in a client/server architecture for receiving information from the Internet;

"a second set of controls having attributes related to a second modality of interaction with a user of the client that being at least one of recognition and audible prompting, . . . the second set of controls being selectively associated with the first set of controls, and the second set of controls being related to client side markup

executable a client browser" – main renderer 14 of a multi-modal presentation manager 11 initiates a second processing thread comprising a speech renderer 16, wherein the speech renderer 16 utilizes a plurality of speech engines 17 for speech recognition and text-to-speech synthesis (column 7, lines 38 to 47: Figure 1); an XML (eXtensible Markup Language) script is implemented in a single authoring format ("defined on an authoring page") (column 5, lines 50 to 56); multi-modal presentation manager 11 controls an application on a web browser or a desktop (column 8, lines 32 to 35: Figure 1); implicitly, a web browser is executed on a client in a client/server architecture for receiving information from the Internet; in deferred rendering and presentation, a speech renderer 16 ("a second set of controls") is "selectively associated with" GUI presentation manager 15 ("a first set of controls") because multi-modal presentation manager 11 automatically integrates and synchronizes voice synthesis and speech recognition functions with the presentation layer of applications (column 6, line 63 to column 7, line 8: Figure 1);

"a module operable on a computer, the module being configured to receive the authoring page, which includes a plurality of the second set of controls, wherein the module is further configured to process the plurality of the second set of controls from the authoring page to generate client side markup from the modality dependent controls that is executable by the client browser on the client in the server/client system in accordance with second set of controls and the attributes of the second set of controls for at least one of recognition and audibly prompting, and wherein the module is configured to use at least one of the first set of controls from the authoring page in order

to generate markup therefrom when processing each of the second set of controls" – main renderer 14 of a multi-modal presentation manager 11 initiates a second processing thread comprising a speech renderer 16, wherein the speech renderer 16 utilizes a plurality of speech engines 17 for speech recognition and text-to-speech synthesis (column 7, lines 38 to 47: Figure 1); an XML (eXtensible Markup Language) script is implemented in a single authoring format ("the authoring page") (column 5, lines 50 to 56); authoring produces content for both GUI presentation manager 15 and speech renderer 16 (column 7, lines 38 to 48).

Concerning independent claims 12, 23, and 52, *Dantzig et al*. discloses grammars in VoiceXML in order to optimize speech recognition functions (column 10, lines 38 to 56), but omits the limitations of "wherein attributes related to recognition include at least one of location of grammar for use in recognition and confidence level thresholds for recognition and wherein attributes related to audible prompting include at least one of inline text for text-to-speech conversion, location of data for audible rendering and playing of a prerecorded audio file". However, *Ladd et al. ('336)* teaches a voice browser for interactive services, where a GRAMMAR input includes a SCR attribute that can be a grammar address (*i.e.*, a URL) for a markup language document: SCR = "gram//.SomeGrammar/month/year" ("location of a grammar for use in recognition"). (Column 20, Line 47 to Column 21, Line 1) Moreover, *Ladd et al. ('336)* provides a voice browser, where a PROMPT element of the markup language is used to define content by <PROMPT> text </PROMPT> that is read by a text-to-speech unit, so that markup of <PROMPT> Please select a soft drink. </PROMPT> includes at least

"inline text for text-to-speech conversion". (Column 16, Line 63 to Column 17, Line 21; Column 18, Lines 33 to 39) An objective is permit users to access information from any location in the world via any suitable network access device. (Column 43, Lines 54 to 63) It would have been obvious to one having ordinary skill in the art to include markup attributes relating to a location of a grammar and inline text for text-to-speech conversion as taught by *Ladd et al. ('336)* in a system and method for generating and presenting multi-modal applications from markup scripts of *Dantzig et al*. for a purpose of permitting users to access information from any location in the world via a suitable network access device.

Concerning independent claim 23, *Dantzig et al*. further discloses "wherein values of the second set of controls are synchronized with the first set of visual controls" – in one aspect, immediate synchronized rendering of the modality-independent document in each of the supported modalities is provided (Abstract); preferably, the multi-modal interface automatically synchronizes I/O events over the plurality of modalities presented (column 2, lines 50 to 53); multi-modal presentation manager 11 provides a runtime environment which integrates and synchronizes a plurality of 'presentation interfaces', enabling I/O events initiated at one 'interface' to be reflected across all interfaces; multi-modal presentation manager 11 provides a mechanism to automatically integrate and synchronize voice synthesis and speech recognition functions with the presentation layer of applications (column 6, line 65 to column 7, line 8: Figure 1).

Concerning claims 14 and 25, *Ladd et al. ('336)* discloses attributes for grammars (column 13, lines 6 to 10), and retrieving grammars from database locations (column 12, lines 7 to 14; column 14, lines 18 to 28) for speech recognition.

Concerning claims 20 to 22 and 31 to 33, *Ladd et al. ('336)* discloses determining an address for playing a prompt to a user (column 13, line 66 to column 14, line 17: Figure 5a: Steps 400, 402, 406); both recorded sound samples (column 15, line 63) and text to speech (TTS) (column 16, lines 11 to 20) are provided.

Concerning claims 13, 15, 24, and 26, *Dantzig et al*. discloses controls relate to grammars for speech recognition (column 9, lines 31 to 39; column 16, lines 6 to 30).

Concerning claims 16 to 17, and 27 to 28, *Dantzig et al*. discloses controls relating to XML (column 5, lines 50 to 56), VoiceXML (a form of XML) (Abstract), and WML (column 6, lines 56 to 62).

Concerning claims 18 to 19, and 29 to 30, *Dantzig et al*. discloses a speech renderer 16 generates audible output by text-to-speech synthesis (column 7, lines 42 to 45).

Concerning claims 34 to 46, *Dantzig et al*. discloses a system and method for generating and presenting multi-modal applications from markup scripts for synchronizing a GUI presentation layer with voice synthesis and speech recognition, but omits details relating to "attributes providing a reference to a location", "a prerecorded audio data file", "an identifier of the associated control", "a question control", "an answer control", "binding the recognition value", and "a confirmation control". However, *Ladd et*

al. ('336) teaches a voice browser for interactive services. An objective is permit users to access information from any location in the world via any suitable network access device. (Column 43, Lines 54 to 63) It would have been obvious to one having ordinary skill in the art to include details disclosed by *Ladd et al. ('336)* in a system and method for generating and presenting multi-modal applications from markup scripts of *Dantzig et al*. for a purpose of permitting users to access information from any location in the world via a suitable network access device.

Concerning claim 34, *Ladd et al. ('336)* discloses a markup language for text to speech; implicitly, when the text is displayed and the speech is produced for an audible prompt, there is an association of attributes between visual controls and audible controls.

Concerning claims 35 to 37, *Ladd et al. ('336)* discloses an option list in a markup language for controlling which choices are available at a network access apparatus (column 28, lines 9 to 60).

Concerning claim 38, *Ladd et al. ('336)* discloses a FORM input to collect an order in response to a prompt, and post the input to an address (column 20, lines 20 to 46); thus, a markup language controls a prompt, then activates an input, and then performs a post operation.

Concerning claims 39 to 43, *Ladd et al. ('336)* discloses a markup language for generating an audible prompt of a question and a grammar for an answer; an answer is followed by, and is activated, a question prompt, where an answer is bound for

recognition by <INPUT TYPE> (column 18, lines 40 to 55); a post operation is "an event related to operation of binding" (column 20, lines 28 to 46).

Concerning claims 44 to 46, *Ladd et al. ('336)* discloses a markup language for re-prompting ("repeating an audible prompt") (column 14, line 57 to column 15, line 16: Figure 5a: Steps 416, 425), and an attribute for confirming a recognition result (column 15, lines 45 to 54: Figure 5a: Step 452).

Claims 47 to 51 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Dantzig et al*. in view of *Ladd et al. ('336)* as applied to claims 23, 39, 40, 45, and 46 above, and further in view of *WCW Working Draft ("Grammar Representation Requirements for Voice Markup Languages")*.

Ladd et al. ('336) discloses a confirmation control to accept an answer as a recognized result that is correct (column 15, lines 44 to 59: Figure 5b: Step 456). Lack of confirmation implicitly denies a recognized result, whereupon the process continues to replay a prompt for a current step so as to correct a recognition result. (Figures 5a and 5b: Step 446) However, *Ladd et al. ('336)* omits an attribute related to a confidence level for confirming, accepting or denying, and correcting a recognition result. *WCW Working Draft* teaches grammars for voice markup languages with attributes, where confidence scoring tightens or relaxes the normal rejection constraints to provide content based control of performance. (Sections 4.3 and 5.1) It would have been obvious to one having ordinary skill in the art to provide confidence scoring as taught by *WCW Working Draft* in the voice browser for interactive services of *Ladd et al. ('336)* for

a purpose of tightening or relaxing rejection constraints to provide content based control of performance.

Response to Arguments

Applicants' arguments filed 07 May 2010 have been fully considered but they are not persuasive.

Basically, Applicants present the same arguments presented previously and add some new ones. Applicants' fundamental position continues to be that the term "modality dependent" is disclosed by their Specification so as to meet the written description requirement under 35 U.S.C. §112, 1st ¶, but that, although the term is literally disclosed by the prior and not by Applicants' Specification, that the prior art somehow fails to disclose the limitation of attributes that are "modality dependent" in the sense contemplated by Applicants. Applicants' argument appears to presuppose that their Specification discloses that their computer code or authoring page or script produces modality dependent interaction in some manner not disclosed by the prior art even though a careful reading of their Specification fails to show any difference in how interaction is produced between a computer and a human for visual and speech modalities as compared with the prior art.

Regarding the rejection for failure to meet the written description requirement under 35 U.S.C. §112, 1st ¶, Applicants argue that an initial burden is not established by the rejection as to why someone of ordinary skill in the art would not find that Applicants' disclosure provides support for the term "modality dependent". However, it is believed

that the rejection meets this threshold requirement of an initial burden because that term is not literally disclosed by the Specification. Applicants then further say that, to comply with the written description requirement, MPEP §2163.02 does not require that the words "modality dependent" be literally disclosed by the Specification. It is agreed that it is true that terms need not be expressly disclosed *in haec verba* to meet the written description requirement if there is some implicit disclosure that would be evident to one skilled in the art. Applicants then say that they have previously submitted numerous reasons why one skilled in the art would know how to interpret the term "modality dependent" just to mean the use of an attribute depends on a specific modality, *i.e.*, a visual modality or a spoken modality. It is agreed, as far as this goes, that someone skilled in the art might have an idea of what modality dependence might mean even though the term is not disclosed by Applicants' Specification. The problem is that Applicants' argument here is disingenuous. To respond to the obviousness rejection, Applicants are then positing that the term "modality dependent" means something else than what they are saying here that one having ordinary skill in the art would understand it to mean. Applicants would have us believe that one having ordinary skill in the art would understand what is meant by "modality dependent" only to later say that modality dependence requires that the two modalities someone be tied into the computer code or authoring page or script in a manner that is not disclosed by the prior art. Effectively, Applicants' later interpretation of "modality dependent" is intended to preclude a modality independent script that, through intermediate processes, is rendered into modality dependent attributes. This is what is not disclosed by Applicants'

Specification. Applicants' claims are directed to "using modality dependent attributes provided from controls on the authoring page". However, Applicants' Specification does not disclose modality dependent controls that are generated from an authoring page, even if it would be reasonable to suppose that one skilled in the art would understand what is meant by modality dependence.

The rejection of claims 1 to 22 for indefiniteness under 35 U.S.C. §112, 2nd ¶, is withdrawn pursuant to Applicants' amendments removing the term "directly" from these claims.

Next, Applicants argue the rejections of the independent claims under 35 U.S.C. §103(a). Concerning independent claim 1, Applicants repeat the argument previously made that "modality dependent attributes" are not disclosed by *Dantzig et al*. One can see immediately then that Applicants' are using a 'point of novelty' argument, where the feature that Applicants allege is not found in the prior art also does not appear to be clearly disclosed by Applicants' own Specification. Again, Applicants argue that because *Dantzig et al*. discloses a modality independent script, then *Dantzig et al*. is incapable of producing modality dependent attributes. However, at the bottom of Page 19 of Applicants' Remarks, Applicants appear to admit that *Dantzig et al*. discloses logic to render an IML script 32 into markup including HTML and VoiceXML. One skilled in art could reasonably understand that if an IML script is used to produce markup including HTML and VoiceXML – *i.e.* literally, "a first and second modality specific representation" as disclosed at Column 2, Lines 64 to 67 of *Dantzig et al*. – then *Dantzig et al*. is disclosing "modality dependent attributes". The question here is, at least

partially, one of 'broadest reasonable interpretation'. When Applicants' Specification is silent on how the term "modality dependent attributes" is to be interpreted, then that term should be construed broadly. Applicants' Specification does not disclose anything that would suggest that the term "modality dependent attributes" are precluded from being produced from a modality independent script. During patent examination, the pending claims must be "given their broadest reasonable interpretation consistent with the specification." *Phillips v. AWH Corp.*, 415 F.3d 1303, 75 USPQ2d 1321 (Fed. Cir. 2005) See MPEP §2111.

Applicants then note an example of an IML script 32 incorporated by *Dantzig et al.* from U.S. Serial No. 09/544,823 at Column 5, Lines 50 to 56. While this application corresponds to U.S. Patent No. 7,685,252 to *Maes et al.*, the portions of the patent cited by Applicants are not particular enlightening. Rather, if Applicants wish to look at *Maes et al.*, then attention should be drawn to Figures 10, 11, 13, and 14, which clearly show how transcoding is performed to produce modality specific GUI rendering and speech rendering. The presence of a transcoder in *Dantzig et al.* and *Maes et al.* cannot be understood to logically imply that these references fail to produce modality dependent attributes under principles of 'broadest reasonable interpretation'.

Applicants then argue that *Dantzig et al.* fails to disclose "a grammar to use for recognition" for independent claim 1, as amended. This is not persuasive. *Dantzig et al.*, in fact, repeatedly discloses a grammar for speech recognition. The clearest exposition of this is at Column 16, Lines 26 to 30, where it is stated that the speech renderer calls a new active grammar for a recognizer. However, *Dantzig et al.* also

says at Column 10, Lines 38 to 40, that VoiceXML makes extensive use of grammars in order to optimize speech recognition functions. More generally, *Dantzig et al*. discloses, at Column 9, Lines 31 to 39, that grammars are generated according to JSGF (Java Speech Grammar Format). The fact that the grammars are associated with speech in JSGF certainly implies that the grammars are associated with an attribute of speech interaction.

Applicants then turn to *Coffman et al*., which is cited in combination with *Dantzig et al*., against independent claim 1. It is maintained that Applicants comments against *Coffman et al*. are not completely clear, although Applicants appears to contend that any modality dependent attributes disclosed by *Coffman et al*. are being read on I/O API 18. However, while it is certainly true that visual attributes of a graphical user interface (GUI) and speech relate to input and output functions, this argument fails to credit that *Coffman et al*. expressly discloses providing modality specific renderings. See Figure 2. Because *Coffman et al*. so expressly discloses modality dependent presentation at least at Column 5, Line 66 to Column 6, Line 3, the current rebuttal is at a loss as to how to reply to Applicants' argument about *Coffman et al*. Applicants do go on to say that there is no mention in *Coffman et al*. of an authoring page, but the modalities of speech and graphics appear to be implemented as VoiceXML and HTML, respectively. (Column 4, Lines 17 to 23) As understood from Applicants' Specification, the term "authoring page" is meant to refer to computer code or a script written in a particular programming language. *Coffman et al*. actually uses the word “authoring” to describe the program for the modalities at Column 20, Line 65 to Column 21, Line 2.

Finally, Applicants say on Page 24 of their Remarks that the Office Action fails to address any of the arguments pointing out the deficiencies from their Response filed 10 September 2009, and this absence needlessly encourages piecemeal prosecution and delays the resolution of issues in a timely manner. This comment is manifestly unfair. Applicants' Response filed 10 September 2009 included only one paragraph on Page 16 traversing the rejection under *Coffman et al*., and the Final Rejection mailed 12 November 2009 included two paragraphs of rebuttal arguments discussing *Coffman et al*. on Pages 18 to 19. Over a long prosecution history, an earnest attempt was made to consider and reply to all of Applicants' arguments, and that appears to be the case in the current instance cited by Applicants, too.

Therefore, the rejections of claims 1 and 3 to 22 under 35 U.S.C. §112, 1st ¶, as failing to comply with the written description requirement; of claims 1 and 4 to 8 under 35 U.S.C. §103(a) as being unpatentable over *Dantzig et al*. in view of *Coffman et al*.; of claims 3 and 9 to 11 under 35 U.S.C. §103(a) as being unpatentable over *Dantzig et al*. in view of *Coffman et al*., and further in view of *Ladd et al. ('336)*; of claims 12 to 46 and 52 under 35 U.S.C. §103(a) as being unpatentable over *Dantzig et al*. in view of *Ladd et al. ('336)*; and of claims 47 to 51 under 35 U.S.C. §103(a) as being unpatentable over *Dantzig et al*. in view of *Ladd et al. ('336)*, and further in view of *WCW Working Draft*, are proper.

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to MARTIN LERNER whose telephone number is (571)272-7608. The examiner can normally be reached on 8:30 AM to 6:00 PM Monday to Thursday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David R. Hudspeth can be reached on (571) 272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Martin Lerner/
Primary Examiner
Art Unit 2626
June 29, 2010